Tugas 1

**ANALISIS PENYAKIT HIPERTENSI**

Dosen pengampu :

*Eka Mala Sari Rochman, S.Kom M.Kom*



**Oleh :**

Daffa Azfaril Naufal (210411100173)

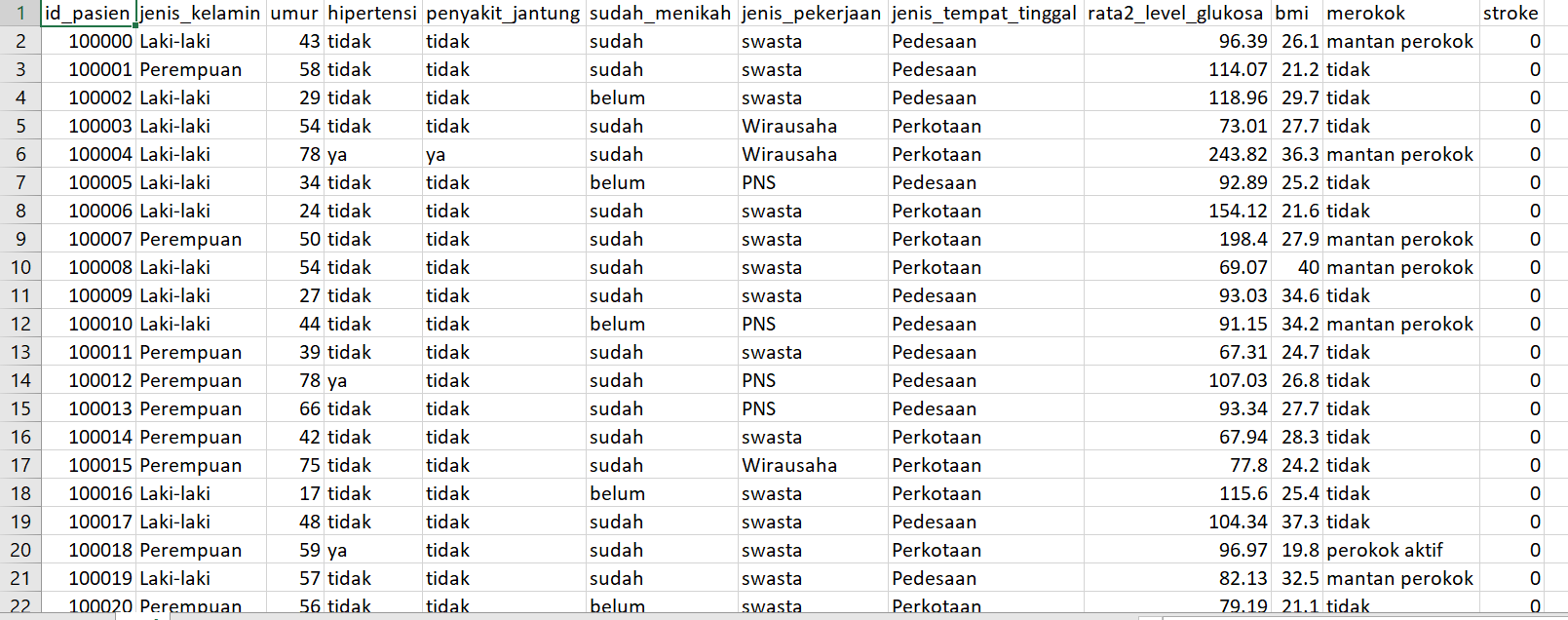
PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS TRUNOJOYO MADURA

**Hipertensi :**

Pada dataset ini terdapat 12 variable yaitu id pasien, jenis kelamin, umur, hipertensi, penyakit jantung, sudah menikah, jenis pekerjaan, jenis tepat tinggal, rata2 level glukosa, bmi, merokok, dan stroke. Dataset ini terdiri dari 28076 data.



Variable targetnya adalah

1. Menderita hipertensi
2. Tidak menderita hipertensi
3. Id pasien

Id pasien digunakan untuk membedakan antara pasien satu dan yang lain, hal ini bertujuan untuk mempermudah mencari data dan juga untuk mengolah data tersebut.

1. Jenis kelamin

Jenis kelamin bertujuan untuk membedakan pasien, dalam hal ini terdapat dua jenis kelamin yaitu perempuan dan laki-laki

1. Umur

Umur merupakan unsur penting karena rata-rata penderita hipertensi menjangkit pasien dengan usia dewasa-tua

1. Hipertensi

Hipertensi adalah suatu kondisi medis di mana tekanan darah di arteri terus meningkat.

1. Penyakit jantung

Penyakit jantung adalah kondisi medis lain yang dikaitkan dengan peningkatan risiko terkena hipertensi

1. Sudah menikah

Pada dataset kebanyakan adalah orang-orang yang sudah menikah

1. Jenis pekerjaan

Jenis pekerjaan pada suatu orang juga dapat menjadi faktor seseorang menderita penyakit hipertensi

1. Jenis tempat tinggal

Pada dataset tersebut terdapat dua tempat tinggal, dan penderita hipertensi tersebar di kedua tempat tersebut

1. Rata2 level glukosa

Rata-rata level glukosa dapat digunakan untuk mengetahui pasien terjangkit hipertensi atau tidak

1. Body Mass Index (bmi)

Orang dengan bmi yang berlebihan juga dapat menyebabkan seseorang untuk menderita penyakit hipertensi

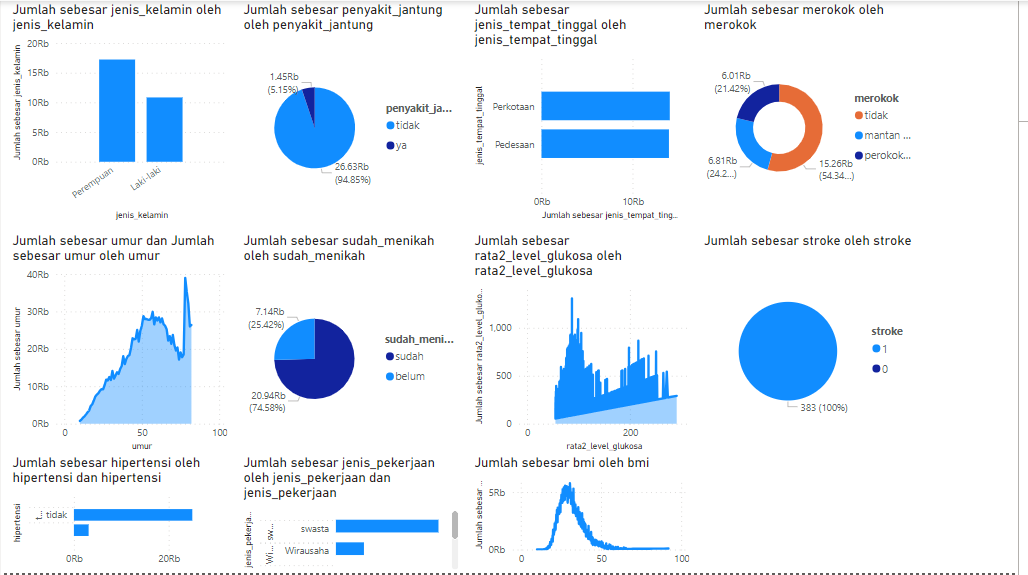
1. Merokok

Riwayat merokok juga dapat menjadi faktor resiko hipertensi dan juga penyakit yang lain. Dalam dataset tersebut terdapat tiga kategori yaitu tidak, mantan perokok, dan perokok aktif

1. Stroke

Stroke adalah penyakit yang biasanya diderita oleh orang-orang dengan usia dewasa-tua

1. **Analisis data menggunakan Power BI**

****

1. **Analisis Parameter**

* Id pasien

Tidak terdapat keterkaitan antara id pasien dengan hipertensi

* Jenis kelamin

Hubungan antara jenis kelamin dengan hipertensi bisa dilihat melalui grade, dimana lebih banyak perempuan yang menderita hipertensi

* Umur

Hubungan antara umur pada grade datasat tersebut umur dewasa-tua rentang memiliki hipertensi

* Hipertensi

Hubungan antara hipertensi dengan grade tersebut adalah menunjukkan apakah pasien tersebut menderita penyakit hipertensi atau tidak

* Penyakit jantung

Hubungan antara penyakit jantung dengan hipertensi adalah jantung yang memompa darah keseluruh tubuh, dapat menyebabkan hipertensi

* Sudah menikah

Hubungan antara sudah menikah dengan grade dataset tersebut adalah orang yang belum menikah memiliki presentase terkena hipertensi lebih besar

* Jenis pekerjaan

Hubungan antara jenis pekerjaan dengan grade ini adalah kebanyakan orang yang bekerja sebagai wirausaha lebih cenderung terkena hipertensi

* Jenis tempat tinggal

Hipertensi dapat di derita oleh orang yang bertempat tinggal di perkotaan dan juga pedesaan

* Rata2 level glukosa

Kadar gula dapat menyebabkan perubahan tekanan darah yang terjadi sehinggan dapat menyebabkan hipertensi

* Bmi

Terdapat keterkaitan antara bmi dengan grade yang terdapat pada dataset ini, dimana semakin tinggi bmi maka semakin rentan terkena hipertensi

1. **Cleaning data**

Terdapat beberapa parameter yang tidak digunakan untuk menentukan penyakit, yaitu :

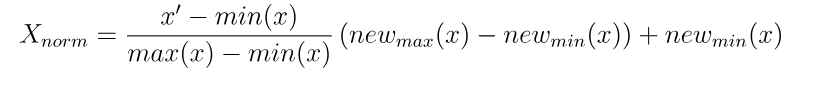
* Id pasien
* Sudah menikah
* Jenis pekerjaan
* Jenis tempat tinggal

1. **Missing data**

Tidak terdapat missing data pada dataset ini

1. **Normalisasi data**

Metode normalisasi min max mengubah kumpulan data menjadi skala 0 (min) hingga 1 (max). Data asli mengalami modifikasi linear dalam prosedur normalisasi data ini. Nilai minimum dan maximum dari data diambil, dan setiap nilai diubah menggunakan rumus di bawah ini :

****

1. **Balancing Data**

Balancing data, dalam konteks analisis data dan machine learning, merujuk pada tindakan untuk menyelaraskan atau mengubah distribusi kelas atau kelompok data agar proporsinya menjadi lebih seimbang. Biasanya, ini diterapkan pada masalah klasifikasi di mana Anda memiliki dua atau lebih kelas atau kelompok, dan distribusi antara kelas-kelas tersebut tidak seimbang.